The complete amino acid sequence of a subunit of the vicilin seed storage protein of Pea (*Pisum sativum* L.)

Hisashi Hirano+, John A. Gatehouse and Donald Boulter*

Department of Botany, University of Durham, Durham DH1 3 LE, UK

Received 18 June 1982

Pea Pisum sativum *Seed* *Storage protein* *Vicilin* *Amino acid sequence*

1. INTRODUCTION

Seeds of *Pisum sativum* L. contain two major storage proteins, legumin and vicilin. Whereas partial amino acid sequence data is available for legumin, not only from the protein [1,2], but also by prediction from a cDNA sequence [3], there is no amino acid sequence data for vicilin. Vicilin in mature seeds has a M_r of 170 000 (7 S) and contains subunits of M_r 50 000, 33 000, 19 000, 15 000, 13 000 and 12 500 [4]. It has been suggested that the subunits of M_r less than 50 000 in vicilin are derived from 50 000 M_r precursor polypeptides by post-translational proteolysis [4]. The present paper describes a purification method for vicilin subunits and presents a complete amino acid sequence for one of the low M_r subunits of vicilin. This subunit has previously been estimated by SDS–PAGE to have a M_r of 13 000; however, the amino acid sequence predicts a M_r of ca. 14 000 and it is referred to as the '14 000 M_r subunit' in this paper. This is the first report of a complete amino acid sequence for any of the subunits of major seed storage proteins.

2. MATERIALS AND METHODS

Vicilin was purified from dry mature seeds (*Pisum sativum* L. cv. Feltham First) (Suttons Seed Ltd.) as described in [4]. The purified vicilin was dissolved in 50 mM Tris–HCl buffer (pH 8.0), containing 8 M urea, and subjected to ion-exchange chromatography on a column of DEAE-cellulose (DE-52, Whatman) (1.6 cm diam., 90 ml) equilibrated with the same buffer. The column was eluted with the starting buffer until all non-bound material had washed off, and was subsequently eluted with a gradient of NaCl (0–0.4 M, 200+200 ml in starting buffer). Eluted fractions were pooled, dialysed against water, lyophilised and analysed by SDS–PAGE. Fractions containing the 14 000 M_r subunit were dissolved in 70% formic acid and loaded onto a column of Sephacryl S-200 (Pharmacia) (1.6 cm diam., 190 ml) equilibrated with 70% formic acid. The column was eluted at 5 ml/h and fractions were collected, dialysed against water and lyophilised. The purified subunit was analysed by SDS–PAGE and two-dimensional PAGE [5].

Portions (2 mg) of the 14 000 subunit were digested with trypsin (Worthington Biochemical) (1:50 enzyme/protein in 0.1 M ammonium bicarbonate, 18 h at 15°C), α-chymotrypsin (Worthington Biochemical) (1:50 enzyme:protein in 0.2 M *N*-ethylmorpholine acetate buffer, pH 8.5, 2 h at 37°C) and *Staphylococcus aureus* V8 protease (Miles Laboratories) (1:30 enzyme:protein in 50 mM ammonium bicarbonate buffer, pH 7.8, 18 h at 37°C). The resulting peptides were dissolved in 100 μl trifluoroacetic acid and separated by HPLC on a reverse-phase column (C18, 10 μm) (Varian). Chromatography conditions were as follows: starting buffer, 0.1% trifluoroacetic acid in water, 5 min; gradient, 0–50% acetonitrile containing 0.1% trifluoroacetic acid over 100 min, 50–

* To whom correspondence should be addressed

+ Permanent address: The Sericultural Experimental Station, M.A.F.F., Tsukuba, Ibaraki, Japan.

70% over 20 min; flow rate 1.0 ml/min; column temperature 30°C.

The amino acid sequences of the peptides were determined by microsequence analysis using 4-*N,N*-dimethylaminoazobenzen-4′-isothiocyanate (DABITC) [6]. C-terminal residue of the subunit was determined by digestion with carboxypeptidase A (Sigma) and the dansyl–Edman procedure [7]. To determine the N-terminal amino acid sequence, the subunit was dissolved into 80 μl of 50% pyridine containing 0.01 M NaOH and subjected to the DABITC analysis.

3. RESULTS AND DISCUSSION

3.1. *Purification of vicilin subunit*

Purified vicilin was dissociated and subjected to ion-exchange chromatography in 8 M urea to effect a partial separation of its component subunits. Of the fractions eluted from the column with the starting buffer the first peak was found to contain a relatively high proportion of the 14 000 M_r subunit by SDS–PAGE. The fraction was further purified by gel filtration on a column of Sephacryl S-200 using 70% formic acid as a denaturing agent. The final peak of protein eluted was found to contain the 14 000 M_r subunit by SDS–PAGE (fig.1). Figure 1 shows that the subunit was at least 90% pure.

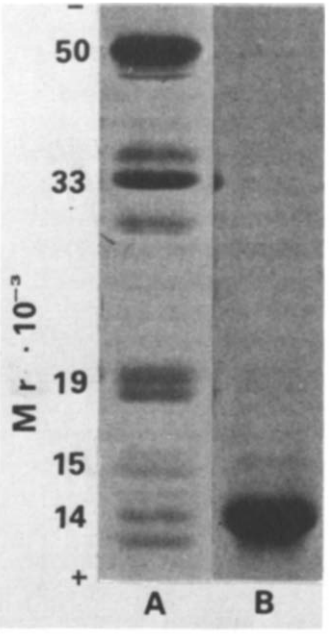

Fig.1. SDS–PAGE of the crude vicilin subunits (A) and the purified 14 000 M_r subunit (B). M_r was estimated using the following proteins as standards: phosphorylase (M_r 97 000), transferrin (M_r 75 000), bovine serum albumin (M_r 67 000), egg albumin (M_r 43 000), chymotrypsinogen (M_r 25 000), myoglobin (M_r 17 200) and cytochrome *c* (M_r 12 500).

3.2. *Amino acid sequence*

The complete amino acid sequence of the 14 000 M_r subunit is shown in fig.2. The subunit contains 124 amino acids which corresponds to a M_r of ca. 14 000. This is in reasonable agreement with the M_r estimated by SDS–PAGE (fig.1).

The three enzymes used to cleave the subunit showed their normal specificities except that the subunit was neither cleaved between positions 28 and 29 with trypsin nor between positions 123 and 124 with *S. aureus* V8 protease. Also, *S. aureus* protease gave an extra cleavage, on the carboxyl side of aspartyl residue in position 71.

The residues in 19 positions are either Leu or Ile which can not be identified in the present study using DABITC analysis. However, the residues in positions 6, 44, 86, 87, 97 and 112 may be Leu since the chymotryptic peptides cleaved on the carboxyl sides of these residues were observed.

Four residue positions in the subunit showed heterogeneity; Arg and Glu in position 18, Ser and Asn in position 39, Glu, Ser and Arg in position 47, and Asp and Asn in position 48. The heterogeneity of the residues in positions 18 and 47 resulted in two tryptic peptides with different N-termini, respectively, and that in position 18, two *Staphylococcus* peptides with different sequences. The 14 000 M_r subunit was separated into six components with different pI by two dimensional PAGE. This difference of pI is possibly due to the heterogeneity observed in these four residues.

There is no sulphur-containing amino acid in the 14 000 M_r subunit, in agreement with the amino acid composition of vicilin [8] and the failure of the 50 000, 33 000 and lower M_r subunits to be cleaved by cyanogen bromide [9].

The amino acid composition determined for the 14 000 M_r subunit was in reasonable agreement with that predicted by the sequence (data not shown).

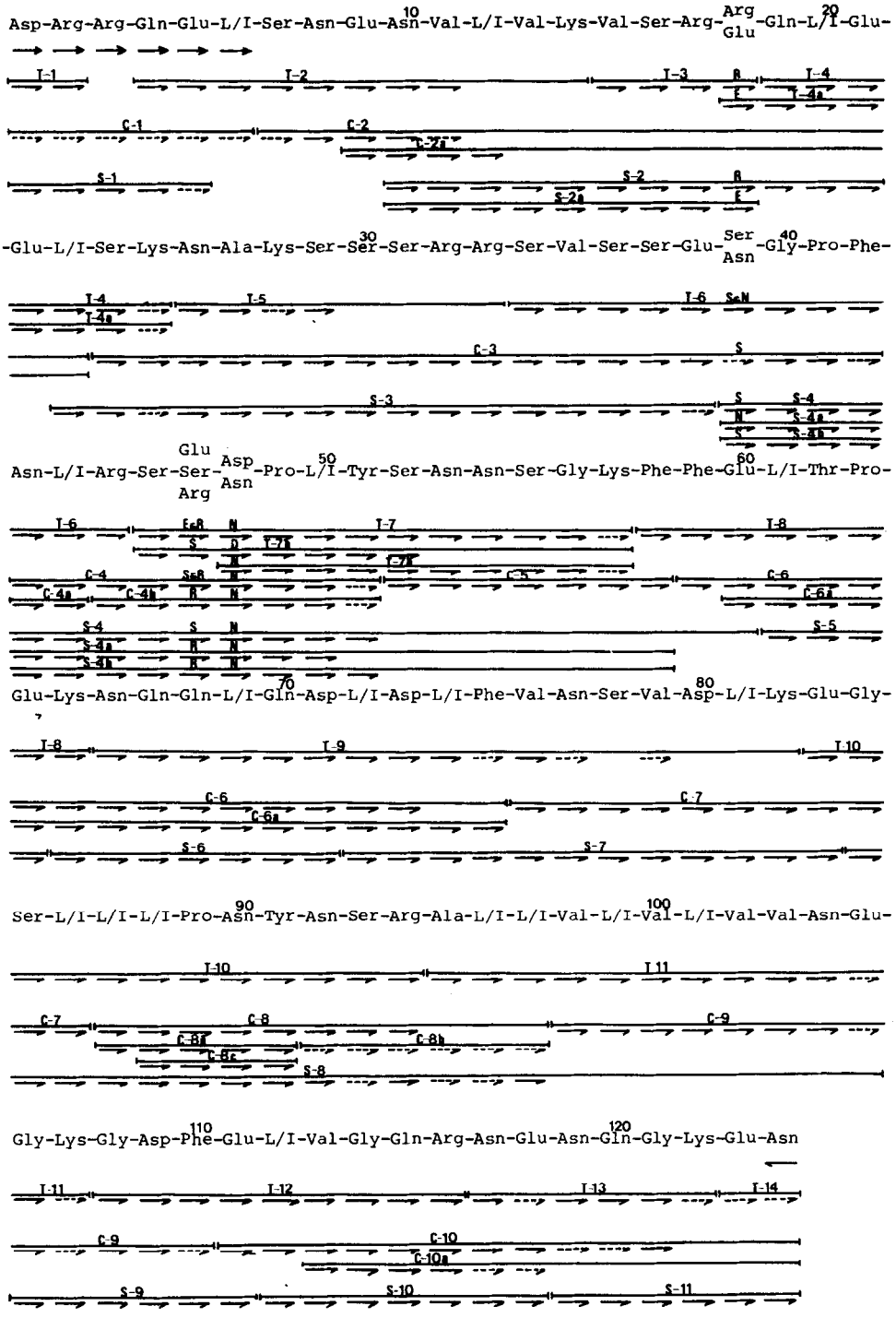

Fig.2. Amino acid sequence of the vicilin 14 000 M_r subunit. T, C and S represent tryptic, chymotryptic and *Staphylococcus* peptides, respectively. Arrow → represents N-terminal amino acid sequence of the subunit. Arrows → and ⇀ were identified by DABITC analysis, and ← by carboxypeptidase A digestion and dansyl–Edman procedure. Residues that were unambiguously identified are indicated with solid arrows, other with broken arrows.

3.3. *Homology with phaseolin*

The amino acid sequence of the vicilin 14 000 M_r subunit is homologous in part to the sequence of phaseolin, the major 7 S storage protein of *Phaseolus vulgaris*, predicted by the determined sequence of a cloned plant gene [10]. If optimal alignments are made, ca. 30% of the amino acid sequence of the vicilin 14 000 M_r subunit is identical to that predicted by nucleotides 411–860 of the phaseolin gene clone. Homologous sequence is found on both sides of an intron in the phaseolin gene. The sequence relation between the 7 S storage proteins in two unrelated legume species supports the suggestion that all legume 7 S storage proteins share a common evolutionary origin [11].

ACKNOWLEDGEMENTS

We wish to thank Mr J. Gilroy for his technical advice of microsequence analysis, Dr D.S. Parker for carrying out amino acid analysis and Drs M. Richardson and G.W. Lycett for their helpful suggestions. H.H. was supported by a fellowship from Science and Technology Agency of Japan.

REFERENCES

[1] Casey, R., March, J.F. and Sanger, E. (1981) Phytochemistry 20, 161–163.
[2] Casey, R., March, J.F., Sharman, J.E. and Short, M.N. (1982) Biochim. Biophys. Acta, in press.
[3] Croy, R.R.D., Lycett, G.W., Gatehouse, J.A., Yarwood, J.N. and Boulter, D. (1982) Nature 295, 76–79.
[4] Gatehouse, J.A., Croy, R.R.D., Morton, H., Tyler, M. and Boulter, D. (1981) Eur. J. Biochem. 118, 627–633.
[5] Hirano, H. (1982) Phytochemistry, in press.
[6] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[7] Richardson, M. (1974) Biochem. J. 137, 101–112.
[8] Croy, R.R.D., Gatehouse, J.A., Tyler, M. and Boulter, D. (1980) Biochem. J. 191, 509–516.
[9] Croy, R.R.D., Gatehouse, J.A., Evans, I.M. and Boulter, D. (1980) Planta 148, 57–63.
[10] Sun, S.M., Slighton, J.L. and Hall, T.C. (1981) Nature 289, 37–41.
[11] Derbyshire, E., Wright, D.J. and Boulter, D. (1976) Phytochemistry 15, 3–24.